

We claim:

1. A method of extracting overlays from video comprising the steps of detecting at least one potential overlay in a video sequence; and verifying that the at least one potential overlay is at least one actual overlay.
2. The method of Claim 1, further comprising the step of: post-processing at least one actual overlay to remove extraneous pixels.
3. The method of Claim 2, wherein said step of post-processing comprises the steps of: computing a variance for each pixel of the at least one actual overlay; and comparing the variance with a threshold to determine whether or not the pixel should be removed as an extraneous pixel.
4. The method of Claim 1, wherein said step of detecting comprises the steps of: performing wavelet decomposition on the video sequence; extracting features based on the results of the wavelet decomposition; and performing neural network processing on the extracted features.
5. The method of Claim 4, wherein said neural network processing step comprises the step of: utilizing three-layer back-propagation neural network processing.
6. The method of Claim 4, wherein said step of verification comprises the steps of: performing temporal verification; and performing spatial verification.
7. The method of Claim 6, wherein said step of temporal verification comprises the steps of: translating said potential overlay over a search range;

for each translated version of said potential overlay, computing a mean square error in a next video frame of said video sequence subsequent to a video frame in which said potential overlay is originally detected;

determining a minimum of the computed mean square errors for said next video frame;

and

comparing the determined minimum mean square error to a threshold.

8. The method of Claim 7, further comprising the steps of:

selecting a particular pixel of said potential overlay and recording its coordinates; and

recording the translated coordinates of said particular pixel corresponding to said determined minimum mean square error.

9. The method of Claim 8, further comprising the step of:

if the determined minimum mean square error does not exceed said threshold,

determining if the coordinates of said particular pixel of said potential overlay match said translated coordinates of said particular pixel corresponding to said determined minimum mean square error.

10. The method of Claim 9, wherein said determining step determines an approximate match.

11. The method of Claim 9, further comprising the step of:

if said determining step determines that there is not a match, performing the sub-steps of:

incrementing an error count; and

comparing said error count to a predetermined threshold; and

if said determining step determines that there is a match, decreasing said error count.

12. The method of Claim 11, wherein said step of decreasing said error count comprises the step of decrementing said error count.
13. The method of Claim 11, wherein said step of decreasing said error count comprises the step of clearing said error count.
14. The method of Claim 11, wherein said steps of computing, determining, recording, and comparing are performed for subsequent video frames of the video sequence as long as said determined minimum mean square error is found not to exceed said threshold and either the coordinates of said particular pixel of said potential overlay match said translated coordinates of said particular pixel corresponding to said determined minimum mean square error or said error count does not exceed said predetermined threshold.
15. The method of Claim 6, wherein said step of performing spatial verification is performed for a candidate overlay determined by said step of performing temporal verification and comprises the steps of:
- determining a structure confidence for said candidate overlay; and
 - determining a texture confidence for said potential overlay.
16. The method of Claim 15, further comprising the steps of:
- determining if said structure confidence meets a first threshold test; and
 - determining if a weighted sum of said structure confidence and said texture confidence meets a second threshold test.
17. The method of Claim 16, wherein said step of said step of determining a texture confidence is performed only if said structure confidence fails to meet said first threshold test.

18. The method of Claim 17, wherein if either of said steps of determining if said structure confidence or said weighted sum meets said respective first or second threshold test is satisfied for the candidate overlay, the candidate overlay is declared to be an overlay; and wherein said steps of determining if said structure and weighted sum meet said respective first and second threshold tests are not satisfied for the candidate overlay, the candidate overlay is determined not to be an overlay.

19. The method of Claim 15, wherein said step of determining a structure confidence comprises the steps of:

analyzing the candidate overlay to determine characters;

analyzing the determined characters for the presence of words; and

setting a numerical value for said structure confidence depending upon the presence of one or more intact words.

20. The method of Claim 19, wherein said step of setting a numerical value comprises the steps of:

setting the structure confidence equal to one if at least one intact word is detected; and

if no intact word is detected, setting the structure confidence equal to a number of correct characters divided by a total number of characters.

21. The method of Claim 15, wherein said step of determining a texture confidence comprises the step of:

setting the texture confidence equal to an average value of outputs of said neural network processing step corresponding to all the pixels in a potential overlay.

22. The method of Claim 1, wherein said step of detecting comprises the step of:

performing template matching to determine the presence of a potential overlay.

23. The method of Claim 22, wherein said step of detecting further comprises the step of:
determining a template to be used in said step of performing template matching.

24. The method of Claim 22, wherein said step of verifying comprises the steps of:
performing frame-to-frame correlation of said potential overlay; and
comparing a result of the frame-to-frame correlation with a threshold to determine if the
potential overlay is an actual overlay or not.

25. The method of Claim 24, wherein said step of performing frame-to-frame correlation
comprises the step of:

forming a mean square error over a set of frames from said video sequence, averaged
over all of the pixels in said potential overlay.

26. A computer-readable medium containing software embodying the method of Claim 1.

27. A computer system comprising:
a computer; and
a computer-readable medium containing software embodying the method of Claim 1.

28. A computer system comprising:
a computer;
a computer-readable medium containing software embodying the method of Claim 4; and
an external processor, in communication with said computer, on which is performed the
step of neural network processing.

29. A method of extracting textual and graphical overlays from video, comprising the steps
of:

detecting at least one potential overlay in a video sequence, said detecting comprising the steps of:

performing wavelet decomposition on the video sequence;
extracting features based on the results of the wavelet decomposition;
performing neural network processing on the extracted features; and
in parallel with said steps of performing wavelet decomposition, extracting features, and performing neural network processing, performing template matching; and verifying that the at least one potential overlay is at least one actual overlay.

30. The method of Claim 29, wherein said step of verifying includes the step of:
performing temporal verification.

31. A method of extracting textual overlays from video, comprising the steps of:
detecting at least one potential overlay in a video sequence, said detecting comprising steps of:

performing wavelet decomposition on the video sequence;
extracting features based on the results of the wavelet decomposition; and
performing neural network processing on the extracted features; and
verifying that the at least one potential overlay is at least one actual overlay.

32. The method of Claim 31, wherein said step of verification comprises the steps of:
performing temporal verification; and
performing spatial verification.

33. The method of Claim 32, wherein said step of spatial verification is performed for a candidate overlay output by said step of temporal verification and comprises the steps of:

determining a structure confidence for said candidate overlay;
determining a layout confidence for said candidate overlay; and
determining a texture confidence for said candidate overlay.

34. The method of Claim 32, wherein said step of performing temporal verification comprises the steps of:

computing a mean square error for each pixel of said potential overlay over a set of video frames of said video sequence;

averaging said mean square error for each pixel over all of the pixels in said potential overlay, thus producing an average mean square error; and

comparing said average mean square error to a threshold to determine if the potential overlay is a candidate overlay or not.

35. A method of extracting graphical overlays from video, comprising the steps of:

detecting at least one potential overlay in a video sequence, said detecting comprising the step of:

performing template matching; and

verifying that the at least one potential overlay is at least one actual overlay, said verifying comprising the step of:

performing frame-to-frame correlation of a potential overlay determined by said detecting step.

36. The method of Claim 35, wherein said step of detecting further comprises the step of:
determining a template to be used in said step of performing template matching.

37. The method of Claim 36, wherein said step of determining a template comprises the step of:

performing addition or frame-by-frame subtraction of video frames.

38. The method of Claim 36, wherein said step of determining a template comprises the steps of:

segmenting video frames into foreground and background objects;

performing correlation tracking to determine if any foreground object remains in the same absolute location in each video frame.